MATCHING IN CLOSED-FORM: EQUILIBRIUM, IDENTIFICATION, AND COMPARATIVE STATICS

RAICHO BOJILOV[§] AND ALFRED GALICHON[†]

Abstract. This paper provides closed-form formulas for a multidimensional two-sided matching problem with transferable utility and heterogeneity in tastes. When the matching surplus is quadratic, the marginal distributions of the characteristics are normal, and when the heterogeneity in tastes is of the continuous logit type, as in Choo and Siow (2006), we show that the optimal matching distribution is also jointly normal and can be computed in closed form from the model primitives. Conversely, the quadratic surplus function can be identified from the optimal matching distribution, also in closed-form. The analytical formulas make it computationally easy to solve problems with even a very large number of matches and allow for quantitative predictions about the evolution of the solution as the technology and the characteristics of the matching populations change.

Keywords: matching, marriage, assignment.

JEL codes: C78, D61, C13.

Date: Date: October 8, 2014. A preliminary version of this paper containing the main results was first presented in September 2012 under the title "Closed-Form Formulas for Multivariate Matching". We would like to thank Arnaud Dupuy, Bernard Salanié and seminar participants at CREST and the 2013 Search and Matching Conference in Paris for helpful comments. Galichon's research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreements no 313699, and from FiME, Laboratoire de Finance des Marchés de l'Energie. Bojilov's work is supported by a grant of the French National Research Agency (ANR), "Investissements d'Avenir" (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

1. Introduction

Many economic problems involve markets with supply and demand restricted to a unit of an indivisible good. Examples include occupational choice, task and schedule assignment, sorting of CEOs to firms, venture capital investment, and the marriage market. Models of discrete choice, selection models such as the famous Roy model, have been applied often to the empirical analysis of such problems. While having important advantages, they generally do not allow to incorporate scarcity constraints and equilibrium effects. Assignment models provide an alternative framework which accommodates these issues. In assignment models, the supply of goods of each type is fixed, and prices adjust at equilibrium so that supply and demand clear. The pioneering work by Choo and Siow [6] proposes an econometric framework for the estimation of assignment models based on the introduction of logit-type unobserved heterogeneity in tastes. Their equilibrium approach has a number of advantages: it incorporates unobserved heterogeneity, allows for matching on many dimensions, and easily lends itself to nonparametric identification. However, the equilibrium quantities in Choo and Siow's setting are defined implicitly by a set of nonlinear equations, and explicit, closed-form solutions, even under further assumptions on the primitives of the model, have been missing so far. The present contribution addresses this issue.

This paper provides closed-form solution to a bipartite matching problem with transferable utility when sorting occurs on multiple dimensions in the presence of unobserved heterogeneity. In particular, it focuses on two related issues: on the one hand, finding and characterizing equilibrium sorting given a surplus technology and, on the other hand, identifying the surplus technology given the optimal matching. Closed-form formulas make it computationally easy to solve problems that involve a large number of matches and characterize the associated equilibrium. They also allow for quantitative predictions about the sorting that will occur and how it evolves as the complementarity on various dimensions changes or as the distribution of characteristics in the populations varies. Moreover, we use the formulas to evaluate the contribution of sorting to social welfare in equilibrium and decompose the impact of changes in structural parameters on individual welfare into selection and price effects. One can also use them as the basis for asymptotic analysis and hypothesis testing.

We follow the broad setting of Choo and Siow [6] by imposing two additional assumptions: quadratic matching surplus and normality on the distributions of the characteristics of the matching parties. We show that the optimal matching distribution is a multivariate Gaussian distribution, and we provide explicit expressions for the relation between the parameters of the matching distribution and the parameters of the primitives of the model. In the case without unobserved heterogeneity, the solution to this problem is well-known in the applied mathematics literature and solved e.g. in [10]; our mathematical contribution here is to move beyond this benchmark and solve the problem with logit heterogeneity. This allows us to solve both the equilibrium characterization problem and the identification problem in closed form. The model assumptions on the functional form and the distributions are quite strong, but they allows us to obtain a very simple closed-form estimator of the surplus and to characterize equilibrium sorting and division of surplus.

This paper is closely related to the literature on identification of the matching function when the surplus is unobservable (see [17] for a good survey), following Choo and Siow [6]. Fox [12] proposes a maximum score estimator which relies on a rank-order property. Galichon and Salanié [15] show that the social welfare in this setting has a tractable formulation involving the entropy of the joint distribution, from which identification can be deduced. Decker et al. [9] provide interesting comparative statics result. Galichon and Salanié ([15], [16]) further introduce a parametric estimator of the surplus function. Dupuy and Galichon [11] extend the model to the continuous case and propose a decomposition of the surplus function into indices of mutual attractiveness, in order to best approximate the matching patterns by lower-dimensional models, and estimate the number of relevant dimensions on which the sorting effectively occurs. Lindenlaub [19] investigates equilibrium properties in multidimensional settings and, independently from the present work, focuses on a matching problem with quadratic surplus and Gaussian distributions of the populations, in the case when there is no unobserved heterogeneity and bivariate observed characteristics. She introduces a very interesting notion of multivariate assortativeness and applies the model to the study of technological change. Finally, in the case of matching with non-transferable utility, Menzel [20] also uses a continuous logit framework which yields a characterization of the equilibrium. As we discuss in the conclusion, the methodology introduced in the present paper could be also applied to his setting.

There is a vast panel of applications of assignment models. In the literature on industrial organization, Fox and Kim [13] used an extension of Choo and Siow's model to study the formation of supply chains. In the corporate governance literature, Gabaix and Landier [14] and Terviö [24] applied an assignment model to explain the sorting of CEOs and firms and the determinants of the CEO compensation. In the marriage literature, Chiappori, Salanié and Weiss [5] used a heteroskedastic version of Choo and Siow's model to estimate the changes in the returns to education on the US marriage market; Chiappori, Oreffice and Quintana-Domeque [4] study the interplay between the sorting on anthropomorphic and on socioeconomic characteristics. In its spirit, our empirical application is closest to Hsieh, Hurst, Jones, and Klenow [18] who use a variation of the Roy model to study discrimination in the US. This list is necessarily very incomplete.

Organization of the paper. The rest of the paper is organized as follows. Section 2 presents the model and states the optimal matching problem. Section 3 starts with the set of conditions that link the solution of the matching problem with the parameters of the social surplus and then derives closed-form expression for the optimal matching and the social surplus. It also characterizes equilibrium compensation. Section 4 concludes.

Notations. Throughout the paper, we denote the transpose of matrix M as M^* and the inverse of this transpose as M^{-*}. Similarly, the inverse of matrix $M^{\frac{1}{2}}$ is $M^{-\frac{1}{2}}$. Let the generalized inverse (see [3]) of a non-square matrix M be denoted $M^+$and its transpose, M^{+*}. All distributions are centered at 0. The random variables are denoted with capital letters, while the corresponding sample points are denoted with small letter.

2. Model

We present the theoretical model in the context of labor matching between a population of firms and a population of workers. We adopt the setting of Dupuy and Galichon [11], extending Choo and Siow [6] to the case when the characteristics of the matching populations are continuous. As it is standard in the search and matching literature, we interpret each firm as a unique job position or task. A worker can be employed by only one firm and a firm can employ only one worker. The populations are assumed to be of equal size, and we assume that it is always better to be matched than to remain alone and generate no surplus.

Each worker has a vector of observable characteristics x known to the econometrician. Similarly, vector y summarizes the observable characteristics of firms. We assume that the surplus enjoyed by worker x choosing occupation y is the sum of three terms: one reflecting her intrinsic taste for occupation y, as predicted by her observable type x; one reflecting the amount of net monetary compensation; and one reflecting unobserved heterogeneity in tastes for occupation y, which is random but has a known distribution conditional on characteristics x. Therefore we maintain:

Assumption 1: Separability. The surplus of worker with characteristics x from matching with firm with characteristics y is

$$\alpha(x,y) + \tau(x,y) + \chi(y) \tag{2.1}$$

where $\alpha(x,y)$ is the nonpecuniary amenity of career with firm y, $\tau(x,y)$ is the monetary transfer or compensation received by the worker from the firm, and $\chi(y)$ is a random utility process drawn by each individual worker and whose distribution is characterized in Assumption 3 below. Similarly, the surplus of firm with characteristics y from matching with worker of characteristics x is

$$\gamma(x,y) - \tau(x,y) + \xi(x) \tag{2.2}$$

where $\gamma(x,y)$ is the economic value created by worker x at firm y, and $\xi(x)$ is a random productivity process whose distribution is also characterized in Assumption 3 below.

Assumption 1 specifies the form of heterogeneity that we investigate, namely heterogeneity in the preferences over the observable characteristics of one's partner. As a result of this assumption, the joint surplus is separable in the sense that $\tilde{\Phi} = \Phi(x,y) + \chi(y) + \xi(x)$, where

$$\Phi(x,y) = \alpha(x,y) + \gamma(x,y)$$

is the matching surplus between worker x and firm y. The first term in the formula for the joint surplus is the observable surplus. The second term allows for unobserved variation in male preference for observed female characteristics, and the third allows for unobserved variation in female characteristics given observed male characteristics. In other words, we rule out an idiosyncratic term that represents the interaction of unobserved male and female characteristics.

Assumption 2: Heterogeneity. The firms and workers observe all of their characteristics but the econometrician observes only the corresponding x and y.

This assumption defines the information structure of the problem: agents have full information, the econometrician does not observe preference shocks.

Assumption 3: Distributions of the unobservables. The functions $\chi(.)$ is modelled as an Extreme-value stochastic process

$$\chi(y) = \max_k \left\{-\lambda(y_k - y) + \sigma_1 \chi_k^i\right\}$$

where $\left(y_k, \chi_k^i\right)$ are the points of a Poisson process on $\mathbb{R}^n \times \mathbb{R}$ of intensity $dy \times e^{-\chi} d\chi$, and $\lambda(z) = 0$ if $z = 0$, $\lambda(z) = +\infty$ otherwise. Similarly,

$$\xi(x) = \max_l \left\{-\lambda(x_l - x) + \sigma_2 \zeta_l^j\right\}$$

where $\left(x_l, \eta_l^j\right)$ are the points of a Poisson process on $\mathbb{R}^n \times \mathbb{R}$ of intensity $dx \times e^{-\varepsilon} d\varepsilon$.

Assumption 3 implies that each worker and each firm draws an infinite, but discrete number of "acquaintances" from the opposite side of the market, along with a random surplus shock. In essence, this specification, it allows for the extension of the convenient logit probability function to the analysis of continuous choice problems. The scale parameter $\sigma = \sigma_1 + \sigma_2$, which captures the total amount of heterogeneity, will play an important role in the sequel.

Assumptions 1-3 are the assumptions of Dupuy and Galichon [11], extending Choo and Siow [6] to the continuous case. We impose two further restrictions, namely normal marginal distributions and quadratic surplus:

Assumption 4: Distribution of the observable. We assume that X has a Gaussian distribution $P = N(0, \Sigma_X)$ and, similarly, that Y has distribution $Q = N(0, \Sigma_Y)$.

Assumption 4 is arguably the strongest restriction of the model. It imposes normal marginal distributions for workers' and firms' characteristics. In some settings, this is a perfectly defendable assumption, while this is not the case in others, due to the presence of discrete characteristics, skewness, fat tails, etc.

Assumption 5: Quadratic surplus. We shall restrict our attention to the very important case when the specification of the surplus function Φ is bilinear, that is

$$\Phi(x, y) = x' A y$$

where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. The $m \times n$ real matrix A is called the *affinity matrix*. We also maintain that if $m \geq n$, the column vectors of A are linearly independent and that if $m < n$, then the row vectors of A are linearly independent.

Assumption 5 imposes quadratic surplus, which is the simplest nontrivial form of the joint surplus yielding complementarities between any pair of firm and worker characteristics. This specification provides an intuitive and meaningful interpretation of the interaction between the characteristics of firms and workers: namely, A_{ij} is simply the strength of the complementarity (positive or negative) between the i-th characteristics of the firm and the j-th characteristics of the worker. Note that x and y can be of different dimension.

We start by reviewing the equilibrium conditions that establish a link between the primitives of the model, particularly the surplus technology, and the optimal matching. As usual, workers maximize utility by solving the following problem:

$$\max_{y} \alpha\left(x,y\right)+\tau\left(x,y\right)+\sigma_1\varepsilon\left(y\right)$$

From the properties of GEV distribution,

$$\log\pi\left(x,y\right)=\frac{\alpha\left(x,y\right)+\tau\left(x,y\right)-a\left(x\right)}{\sigma_1} \tag{2.3}$$

where $a\left(x\right)$ is a normalization term that depends only on x, not on y. Similarly, one obtains another condition that links the optimal matching distribution and the technology of the firm $\gamma\left(x,y\right)$

$$\log\pi\left(x,y\right)=\frac{\gamma\left(x,y\right)-\tau\left(x,y\right)-b\left(y\right)}{\sigma_2} \tag{2.4}$$

Using the equilibrium conditions, we solve for $\tau\left(x,y\right)$ and $\pi\left(x,y\right)$ by combining (2.4) and (2.3) to obtains:

$$\log\pi\left(x,y\right) = \frac{\Phi\left(x,y\right)-a\left(x\right)-b\left(y\right)}{\sigma} \tag{2.5}$$

$$\tau\left(x,y\right) = \frac{\sigma_1\left(\gamma\left(x,y\right)-b\left(y\right)\right)-\sigma_2\left(\alpha\left(x,y\right)-a\left(x\right)\right)}{\sigma} \tag{2.6}$$

where $\sigma=\sigma_1+\sigma_2$.

We recall the characterization of the equilibrium given in Dupuy and Galichon [11], who show how the problem of estimating the equilibrium allocation, or optimal matching, can be solved by estimating the corresponding social welfare problem of the planner. Given Assumptions 1 to 5, a matching assignment is a distribution π over the characteristics of firms and workers (X,Y) with marginals for X and Y equal to P and Q respectively. Let $\mathcal{M}\left(P,Q\right)$ be the set of matching assignments, i.e. distributions over $\left(X,Y\right)$, such that $X\sim P$, and $Y\sim Q$. Theorem 1 in [11] implies that:

(i) At equilibrium, the optimal matching $\pi\in M\left(P,Q\right)$ is the unique maximizer of the social welfare

$$\mathcal{W}\left(A\right)=\sup_{\pi\in\mathcal{M}(P,Q)}\left(\mathbb{E}_\pi\left[X'AY\right]-\sigma\mathcal{E}\left(\pi\right)\right) \tag{2.7}$$

where

$$\begin{aligned}\mathcal{E}(\pi) &= \mathbb{E}_\pi\left[\ln \pi(X,Y)\right] \text{ if } \pi \text{ is absolutely continuous}\\ &= +\infty \text{ otherwise.}\end{aligned}$$

(ii) π_{XY} is solution of (2.7) if and only if it is solution of

$$\begin{cases} A_{ij} = \sigma \frac{\partial \log \pi_{Y|X}(y|x)}{\partial x_i \partial y_j} \\ \pi_{XY} \in \mathcal{M}(P,Q) \end{cases}. \tag{2.8}$$

(iii) for any $\sigma > 0$, the solution of (2.8) exists and is unique.

These results can be seen as an extension of the Monge-Kantorovich theory (see Chapter 2 of [25]). They imply that the introduction of unobserved heterogeneity naturally leads to the classical matching problem with an additional information term that attracts the optimal solution toward a random matching: the entropy of the joint distribution $\mathbb{E}_\pi\left[\ln \pi(X,Y)\right]$. If σ is large, then optimality requires minimizing the mutual information which happens when firms and workers are matched randomly to each other. On the other hand, if σ is small, optimality requires maximizing the observable surplus.

3. Closed-Form Formulas

This section contains our results. We start by discussing how the quadratic cost setting and the distributional assumptions give rise to a tractable condition that allows us to find a closed-form formula for the optimal matching in terms of the affinity matrix A, and the parameters that describe the distributions of X and Y. Then, we use the same condition to recover the expression of the surplus, namely the affinity matrix A, in terms of the observed data. Denote

$$\Sigma_{XY} = \left(\mathbb{E}_\pi\left[X_i Y_j\right]\right)_{ij} = \mathbb{E}_\pi\left[XY'\right] \tag{3.1}$$

to be the *cross-covariance matrix* computed at the optimal π solution of (2.7). We will show that under assumptions 1 to 5 the optimal solution is normal and that $(\Sigma_X, \Sigma_Y, \Sigma_{XY})$ completely parameterize the distribution π_{XY}.

We consider two related problems. The **equilibrium computation problem** requires, given a surplus technology, to determine the optimal matching π. Finding out how to do this in a tractable way, possibly in closed form, allows to make quantitative predictions about the sorting that will occur on the market and to derive comparative statics. In contrast, the **identification problem** starts with an observed matching that is assumed to be stable (or equivalently, optimal) and the goal is to determine the underlying surplus technology. This problem is the inverse problem of the former and is of primary interest to the econometrician. The conditions (2.8), in particular the first one, play a crucial role in our analysis because establish a link between the cross-covariance matrix and the affinity matrix. Thus, the equilibrium computation problem involves solving the first-order conditions (2.8) for the optimal matching and Σ_{XY}, while the identification problem involves solving the same conditions but for A given Σ_{XY}. Closed-form formulas shall be provided for these two problems.

3.1. **Equilibrium Computation Problem.** The following result provides an explicit solution to the equilibrium computation problem. Recall that, for a given matrix M, M^{+} denotes the generalized inverse, which coincides with M^{-1} when M is invertible.

Theorem 1. *Assume $\sigma > 0$ and suppose that Assumptions 1 to 5 hold. Let $(X, Y) \sim \pi$ be the solution of (2.7). Then:*

(i) The relation between X and Y takes the following form

$$Y = TX + \epsilon \tag{3.2}$$

where $\epsilon \sim N(0, V)$ is a random vector independent from X, and matrices T and V are given by

$$T = \Delta \left(\Delta A^{*} \Sigma_X A \Delta \right)^{-1/2} \Delta A^{*} - \frac{\sigma}{2} A^{+} \Sigma_X^{-1} \tag{3.3}$$

$$V = \sigma \Delta \left(\Delta A^{*} \Sigma_X A \Delta \right)^{-1/2} \Delta - \frac{\sigma^2}{2} A^{+} \Sigma_X^{-1} A^{+*} \tag{3.4}$$

where matrix Δ in turn expresses as

$$\Delta = \left(\frac{\sigma^2}{4} A^{+} \Sigma_X^{-1} A^{+*} + \Sigma_Y \right)^{1/2}.$$

(ii) The optimal matching π is the Gaussian distribution $N(0,\Sigma)$ where

$$\Sigma = \begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{XY}^* & \Sigma_Y \end{pmatrix}, \tag{3.5}$$

and the cross-covariance matrix of X and Y, namely $\Sigma_{XY} = \mathbb{E}_\pi [XY']$ is given by

$$\Sigma_{XY} = \Sigma_X A\Delta \left(\Delta A^* \Sigma_X A\Delta\right)^{-1/2} \Delta - \frac{\sigma}{2} A^{+*}. \tag{3.6}$$

Proof. Under Assumptions 1-5, and as recalled in Section 2, the solution to the matching assignment problem (2.7) is unique and characterized by conditions (2.8). Hence it is sufficient just to verify that taking π as the p.d.f. of the $N(0,\Sigma)$ distribution where Σ satisfies (3.5) and (3.6) satisfies conditions (2.8) characterizing the optimal assignment.

Verification of optimality. Let π be the distribution of (X,Y) where $X \sim N(0,\Sigma_X)$ and the distribution of Y conditional on X is given by (3.16), for T and V given by (3.3) and (3.4). We verify that π satisfies the two conditions of characterization (2.8). Indeed, one has

$$\partial^2_{xy} \log \pi_{Y|X}(y|x) = T^* V^{-1} = A$$

and

$$\begin{aligned} var(Y) &= T\Sigma_X T^* + V \\ &= \left(\Delta \left(\Delta A^* \Sigma_X A\Delta\right)^{-1/2} \Delta A^* - \frac{\sigma}{2} A^+ \Sigma_X^{-1}\right) \Sigma_X T^* + V \\ &= \Delta^2 - \sigma\Delta \left(\Delta A^* \Sigma_X A\Delta\right)^{-1/2} \Delta + \frac{\sigma^2}{4} A^+ \Sigma_X^{-1} A^{+*} + V \\ &= \Delta^2 - \frac{\sigma^2}{4} A^+ \Sigma_X^{-1} A^{+*} = \Sigma_Y. \end{aligned}$$

Hence, condition (2.8) is verified and π is optimal for (2.7), QED.

Although this proof by verification is sufficient from a pure mathematical point of view, it is not very didactic as it is not informative about how the formula for Σ_{XY} was obtained. For the convenience of the reader, we shall now flesh out how to solve for Σ_{XY}.

Solving for Σ_{XY}. In this paragraph, we look for a necessary condition on Σ so that $N(0,\Sigma)$ be a solution of (2.7). The derivation made before Theorem (2) implies the following relation to be inverted between A and Σ_{XY}

$$\frac{A}{\sigma} = \left(\Sigma_Y \left(\Sigma_{XY}\right)^{+} \Sigma_X - \Sigma_{XY}^{*}\right)^{+}.$$

Some algebra leads to the following quadratic equation in Σ_{XY}

$$0 = \Sigma_{XY}^{*}\Sigma_X^{-1}\Sigma_{XY} + \sigma A^{+}\Sigma_X^{-1}\Sigma_{XY} - \Sigma_Y \tag{3.7}$$

Let

$$\Psi = \Sigma_X^{-1/2}\Sigma_{XY} \tag{3.8}$$

$$B = \frac{\sigma}{2}A^{+}\Sigma_X^{-1/2} \tag{3.9}$$

$$C = -\Sigma_Y \tag{3.10}$$

so that equation (3.7) becomes

$$\Psi^{*}\Psi + 2B\Psi + C = 0 \tag{3.11}$$

hence

$$2B\Psi = \Sigma_Y - \Sigma_{XY}^{*}\Sigma_X^{-1}\Sigma_{XY}$$

is the variance-covariance matrix of the residual ϵ of the regression of Y on X

$$\Sigma_Y = T\Sigma_X T^{*} + V \tag{3.12}$$

and T and V are given by expressions:

$$T = \Sigma_{XY}^{*}\Sigma_X^{-1} \tag{3.13}$$

$$V = \Sigma_Y - \Sigma_{XY}^{*}\Sigma_X^{-1}\Sigma_{XY} \tag{3.14}$$

Hence, $B\Psi$ is symmetric positive. Equation (3.11) can be rewritten as

$$\left(\Psi + B^{*}\right)^{*}\left(\Psi + B^{*}\right) = BB^{*} - C$$

Thus, we have for the solution that

$$\Psi = U\Delta - B^{*} \tag{3.15}$$

where

$$\Delta = (BB^* - C)^{1/2} = \left(\frac{\sigma^2}{4} A^+ \Sigma_X^{-1} A^{+*} + \Sigma_Y \right)^{1/2}$$

is a positive semi-definite matrix and U is an orthogonal matrix to be defined. Note that $BB^* - C = BB^* + \Sigma_Y$ is a positive symmetric matrix so its square root exists and is uniquely defined.

We now determine U. Since $B\Psi$ is symmetric positive, $BU\Delta = B\Psi + BB^*$ is symmetric positive, hence $\Delta^{-1}BU = \Delta^{-1}(BU\Delta)\Delta^{-1}$ as Δ is symmetric and invertible. Let

$$\Gamma = \Delta^{-1}B$$

Since ΓU is symmetric positive, $(\Gamma U)^2 = \Gamma U U^* \Gamma^* = \Gamma\Gamma^*$, which implies that $\Gamma U = (\Gamma\Gamma^*)^{1/2}$. Hence U is determined by

$$U = \Gamma^{-1} (\Gamma\Gamma^*)^{1/2} = B^{-1}\Delta \left(\Delta^{-1}BB^*\Delta^{-1}\right)^{1/2}.$$

Substituting in (3.15) yields

$$\Psi = U\Delta - B^* = B^{-1}\Delta \left(\Delta^{-1}BB^*\Delta^{-1}\right)^{1/2} \Delta - B^*$$

and replacing B, Σ_{XY}, T and V by their respective expressions (3.9), (3.8), (3.13) and (3.14) gives

$$\begin{aligned}
\Psi &= \Sigma_X^{1/2} A\Delta \left(\Delta^{-1}\left(A^+\Sigma_X^{-1}A^{+*}\right)\Delta^{-1}\right)^{1/2} \Delta - \frac{1}{2}\Sigma_X^{-1/2}A^{+*} \\
\Sigma_{XY} &= \Sigma_X A\Delta \left(\Delta^{-1}\left(A^+\Sigma_X^{-1}A^{+*}\right)\Delta^{-1}\right)^{1/2} \Delta - \frac{\sigma}{2}A^{+*} \\
T &= \Delta \left(\Delta A^*\Sigma_X^{-1}A\Delta\right)^{-1/2} \Delta A^* - \frac{\sigma}{2}A^+\Sigma_X^{-1} \\
V &= \sigma\Delta \left(\Delta^{-1}\left(A^+\Sigma_X^{-1}A^{+*}\right)\Delta^{-1}\right)^{1/2} \Delta - \frac{\sigma^2}{2}A^+\Sigma_x^{-1}A^{+*}.
\end{aligned}$$

∎

Hence it turns out that the matching π optimal solution of (2.7) is jointly normal too, a fact that was not a priori obvious, as π may have normal marginal distributions for X and Y without necessarily being jointly normal. We first consider how the solution evolves both when the unobserved heterogeneity declines to zero and when it increases to infinity.

The following lemma provides these limiting results, recovering well-known results in the optimal transportation literature (see e.g. [10]).

Lemma 1. *Suppose that Assumptions 1 to 5 hold, and let Σ_{XY} be the cross-covariance matrix under the optimal matching. Then:*

(i). $\lim_{\sigma\to 0}\Sigma_{XY} = \Sigma_X A\Sigma_Y^{1/2}\left(\Sigma_Y^{1/2}A^*\Sigma_X A\Sigma_Y^{1/2}\right)^{-1/2}\Sigma_Y^{1/2}$

(ii). $\lim_{\sigma\to\infty}\Sigma_{XY} = 0$.

When the importance of unobserved variables increases, the optimal matching ceases to depend on the observable characteristics X and Y, so in the limit the cross-covariance matrix converges to zero. In the other limit case, when there is no unobserved heterogeneity, the optimal matching problem becomes a special case of Brenier's Theorem (see [25], Ch. 2) and there exists an optimal matching map given by $Y = T\left(A^*X\right)$ where T is the gradient of a convex function. In the likely presence of heterogeneity, however, there exists no such optimal map which complicates the theoretical and econometric problems. Nevertheless, below we show that it is still possible to establish a "regression-style" relation between X and Y in our setting.

3.2. **Identification Problem.** The identification problem is simpler. From Theorem 1, (X, Y) is jointly normal, so one may regress Y on X and write

$$Y = TX + \epsilon \tag{3.16}$$

where $\epsilon \sim N\left(0, V\right)$ is independent from X. Consequently, we can express $\Sigma_Y = \mathbb{E}\left[YY'\right]$ and $\Sigma^*_{XY} = \mathbb{E}\left[YX'\right]$ as

$$\Sigma_Y = T\Sigma_X T^* + V$$

$$\Sigma^*_{XY} = T\Sigma_X$$

and, solving for T and V, **we arrive again at expressions** (3.13) and (3.14):

$$\begin{aligned} T &= \Sigma^*_{XY}\Sigma_X^{-1} \\ V &= \Sigma_Y - \Sigma^*_{XY}\Sigma_X^{-1}\Sigma_{XY} \end{aligned}$$

Since the distribution of Y conditional on X, $\pi_{Y|X}(y|x)$, is also normal,

$$\partial^2_{xy} \log \pi_{Y|X}(y|x) = T^* V^{-1} = \Sigma_X^{-1} \Sigma_{XY} \left(\Sigma_Y - \Sigma^*_{XY} \Sigma_X^{-1} \Sigma_{XY} \right)^{-1}.$$

Introducing this expression in condition (2.8) implies

$$\begin{aligned} A &= \sigma \partial^2_{xy} \log \pi_{Y|X}(y|x) \\ &= \sigma \Sigma_X^{-1} \Sigma_{XY} \left(\Sigma_Y - \Sigma^*_{XY} \Sigma_X^{-1} \Sigma_{XY} \right)^{-1} \end{aligned}$$

This expression can be further simplified to

$$A = \left(\Sigma_Y \Sigma^+_{XY} \Sigma_X - \Sigma^*_{XY} \right)^+$$

As in all choice models, the model is identified up to a location and scale parameter: the identified parameter is the rescaled affinity matrix A/σ. Consequently, we normalize the scale of heterogeneity to one, $\sigma = 1$, and estimate the norm of the affinity matrix. If there are compensation data, one can recover also $\alpha(x, y)$ and $\gamma(x, y)$,as well as the ratio of the scale parameters σ_1 and σ_2, using condition (2.6) which links compensation and the utility and profit functions. We formalize the identification result for the affinity matrix A in the theorem below, whose proof is immediate given the proof of Theorem 1.

Theorem 2. *Suppose that Assumptions 1 to 5 hold and let $\sigma = 1$. Then the affinity matrix A is given by*

$$\begin{aligned} A &= \Sigma_X^{-1} \Sigma_{XY} \left(\Sigma_Y - \Sigma^*_{XY} \Sigma_X^{-1} \Sigma_{XY} \right)^{-1} \\ &= \left(\Sigma_Y \Sigma^+_{XY} \Sigma_X - \Sigma^*_{XY} \right)^+ . \end{aligned} \tag{3.17}$$

3.3. **Comparative Statics.** In many problems, it is important to compute how an increase in the complementarity between two characteristics in the surplus formula affects the optimal matching. Similarly, one may want to find how the optimal matching changes as the variances in the matching populations increase. From an econometric point of view, the differentiation of the estimator of the affinity matrix with respect to the summary statistics will allow us to derive central limit theorems and to compute confidence intervals. In this spirit, Decker et al. [9] provide interesting comparative statics, although their formula are not explicit. Our formulas allow to provide simple, closed-form formulas for these comparative statics, with the help of some background knowledge on the Kronecker product and matrix differentiation, for which we now give the basic definitions (a more elaborate tutorial is given in the Appendix). As defined in that appendix, $\mathbb{T}$ is the operator such that

$$\mathbb{T}vec\left(M\right)=vec\left(M^{*}\right).$$

The following theorem summarizes our results on comparative statics.

Theorem 3. *Suppose that Assumptions 1 to 5 hold. Then:*

(i) *The rate of change of the matching estimator with respect to the covariation between the matching populations (keeping Σ_X and Σ_Y constant) is*

$$\frac{\partial A}{\partial\Sigma_{XY}}=\left(A^{*}\otimes A\right)\left[\left(\Sigma_{X}\Sigma_{XY}^{+*}\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right)+\mathbb{T}\right],\tag{3.18}$$

and the rates of change with respect to the variances of the matching populations (keeping Σ_{XY} constant) are

$$\begin{aligned}\frac{\partial A}{\partial\Sigma_{X}} &= -\left(A^{*}\otimes A\right)\left[I\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right],&(3.19)\\ \frac{\partial A}{\partial\Sigma_{Y}} &= -\left(A^{*}\otimes A\right)\left[\Sigma_{X}\Sigma_{XY}^{+*}\otimes I\right].&(3.20)\end{aligned}$$

(ii) *The rate of change of the optimal cross-covariance matrix with respect to the affinity matrix A (keeping Σ_X and Σ_Y constant) is*

$$\frac{\partial\Sigma_{XY}}{\partial A}=\left[\left(\Sigma_{X}\Sigma_{XY}^{+*}\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(A^{+*}\otimes A^{+}\right),\tag{3.21}$$

and the rates of change with respect to the variances of the matching populations (keeping A constant) are

$$\begin{aligned}\frac{\partial\Sigma_{XY}}{\partial\Sigma_{X}} &= \left[\left(\Sigma_{X}\Sigma_{XY}^{+*}\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(I\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right),&(3.22)\\ \frac{\partial\Sigma_{XY}}{\partial\Sigma_{Y}} &= \left[\left(\Sigma_{X}\Sigma_{XY}^{+*}\otimes\Sigma_{Y}\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(\Sigma_{X}\Sigma_{XY}^{+*}\otimes I\right).&(3.23)\end{aligned}$$

Proof of Theorem 3. We use extensively the properties of the Kronecker product and the *vec* operator recalled in the Appendix.

We start with the comparative statics for A. By differentiation of Equation (3.17),

$$\begin{aligned}\frac{\partial A}{\partial \Sigma_{XY}} &= -\left(A^{*}\otimes A\right)\left[-\left(\Sigma_X\otimes\Sigma_Y\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)-\mathbb{T}\right]\\ &= \left(A^{*}\otimes A\right)\left[\left(\Sigma_X\Sigma_{XY}^{+*}\otimes\Sigma_Y\Sigma_{XY}^{+}\right)+\mathbb{T}\right]\end{aligned}$$

Similarly, the derivatives with respect to Σ_X and Σ_Y are:

$$\frac{\partial A}{\partial \Sigma_X} = -\left(A^{*}\otimes A\right)\left[I\otimes\Sigma_Y\Sigma_{XY}^{+}\right]$$

$$\frac{\partial A}{\partial \Sigma_Y} = -\left(A^{*}\otimes A\right)\left[\Sigma_X\Sigma_{XY}^{+*}\otimes I\right]$$

We can also use equation (3.17) to derive implicitly $\frac{\partial \Sigma_{XY}}{\partial A}$, $\frac{\partial \Sigma_{XY}}{\partial \Sigma_X}$, and $\frac{\partial \Sigma_{XY}}{\partial \Sigma_Y}$. By differentiation of the two sides with respect to A, one gets

$$-\left(A^{+*}\otimes A^{+}\right) = -\left(\Sigma_X\otimes\Sigma_Y\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)\frac{\partial \Sigma_{XY}}{\partial A}-\mathbb{T}\frac{\partial \Sigma_{XY}}{\partial A}$$

$$\left[\left(\Sigma_X\otimes\Sigma_Y\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)+\mathbb{T}\right]\frac{\partial \Sigma_{XY}}{\partial A} = \left(A^{+*}\otimes A^{+}\right)$$

$$\begin{aligned}\frac{\partial \Sigma_{XY}}{\partial A} &= \left[\left(\Sigma_X\otimes\Sigma_Y\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(A^{+*}\otimes A^{+}\right)\\ &= \left[\left(\Sigma_X\Sigma_{XY}^{+*}\otimes\Sigma_Y\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{-1}\left(A^{+*}\otimes A^{+}\right)\end{aligned}$$

Similarly, by differentiation with respect to Σ_X and Σ_Y, one gets

$$\begin{aligned}\frac{\partial \Sigma_{XY}}{\partial \Sigma_X} &= \left[\mathbb{T}+\left(\Sigma_X\otimes I\right)\left(I\otimes\Sigma_Y\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)\right]^{+}\left(I\otimes\Sigma_Y\Sigma_{XY}^{+}\right)\\ &= \left[\left(\Sigma_X\Sigma_{XY}^{+*}\otimes\Sigma_Y\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(I\otimes\Sigma_Y\Sigma_{XY}^{+}\right)\\ \frac{\partial \Sigma_{XY}}{\partial \Sigma_Y} &= \left[\left(I\otimes\Sigma_Y\right)\left(\Sigma_X\otimes I\right)\left(\Sigma_{XY}^{+*}\otimes\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(\Sigma_X\Sigma_{XY}^{+*}\otimes I\right)\\ &= \left[\left(\Sigma_X\Sigma_{XY}^{+*}\otimes\Sigma_Y\Sigma_{XY}^{+}\right)+\mathbb{T}\right]^{+}\left(\Sigma_X\Sigma_{XY}^{+*}\otimes I\right).\end{aligned}$$

∎

This theorem can be used to evaluate counterfactual experiments. A change in technology, for instance, may affect the affinity matrix A. It can also be used to derive asymptotic properties, using a standard delta method and the expression of the derivatives (3.18)-(3.20).

4. Conclusion

This paper adopts the framework of Choo and Siow [6] and Galichon and Salanié [15] and imposes additional functional form and distributional assumptions that allow us to derive closed-form formulas for the surplus and the optimal matching. In this setting, one may solve easily problems that in the past have been computationally intensive. Moreover, the results allow for the characterization of the optimal solution and how it changes as the primitives of the model change. The model in this paper is suited to static matching problems with exogenously fixed populations in which all participants are matched. Thus, it covers a number of contexts with potential for future applied work, such as matching of workers with tasks, scheduling, matching between specialists and firms or top managers and firms.

While our framework is Transferable Utility (TU), this paper's main argument may be carried without difficulty into some instances of the Nontransferable Utility (NTU) framework. Indeed, a recent contribution by Menzel (2014) implies that in the case with no singles and a particular structure of (logit) heterogeneity, the equilibrium matching in the TU and the NTU cases coincide. Hence, the results of the present paper would apply to that framework as well.

Appendix: Matrix Differentiation

The Kronecker product and the vectorization operation are extremely useful when it comes to studying asymptotic properties involving matrices. The idea is that matrices $m \times n$ can be seen as vectors in $\mathbb{R}^{mn}$, and linear operation on such matrices can be seen as higher order matrices. To do this, the fundamental tool is the *vectorization* operation. Introduce τ_{mn} a collection of invertible maps from $\{1, ..., m\} \times \{1, ..., n\}$ onto $\{1, ..., mn\}$. For instance, $\tau_{mn}(i, j) = j(n-1) + i$ will do.

Definition 1. *For (M) a $m \times n$ matrix, $vec(M)$ is the vector $v \in \mathbb{R}^{mn}$ such that $v_{\tau_{mn}(i,j)} = M_{ij}$.*

Next, we introduce the transposition tensor $\mathbb{T}_{m,n}$ as the $mn \times mn$ matrix such that

$$\mathbb{T}_{m,n} vec\left(M\right) = vec\left(M^*\right).$$

Equivalently, the r, s entry of $\mathbb{T}_{m,n}$ is one if $\tau_{mn}\left(r\right) = \left(j, i\right)$ where $\left(i, j\right) = \tau_{mn}\left(s\right)$, zero otherwise. Note that $\mathbb{T}_{m,n} = \mathbb{T}_{n,m}^{-1}$ and that $\mathbb{T}_{m,n} = \mathbb{T}_{n,m}^{*}$. The next definition deals with Kronecker product, which is closely related to vectorization.

Definition 2. *Let A be a $m \times p$ matrix and B an $n \times q$ matrix. One defines the Kronecker product $A \otimes B$ as the $mn \times pq$ matrix such that*

$$\left(A \otimes B\right)_{\tau_{mn}(i,k)\tau_{pq}(j,l)} = A_{ij} B_{kl}.$$

The following fundamental property characterizes the Kronecker product.

Fact 1. *For all $q \times p$ matrix X,*

$$vec\left(BXA^T\right) = \left(A \otimes B\right) vec\left(X\right)$$

The following important basic properties follow.

Fact 2. *Let A be a $m \times p$ matrix and B an $n \times q$ matrix. Then:*

1. *(Associativity)* $\left(A \otimes B\right) \otimes C = A \otimes \left(B \otimes C\right)$.
2. *(Distributivity)* $A \otimes \left(B + C\right) = A \otimes B + A \otimes C$.
3. *(Multilinearity) For λ and μ scalars,* $\lambda A \otimes \mu B = \lambda\mu \left(A \otimes B\right)$
4. *For matrices of appropriate size,* $\left(A \otimes B\right)\left(C \otimes D\right) = \left(AC\right) \otimes \left(BD\right)$.
5. $\left(A \otimes B\right)^* = A^* \otimes B^*$.
6. *If A and B are invertible,* $\left(A \otimes B\right)^{-1} = A^{-1} \otimes B^{-1}$.
7. *For vectors a and b, $a' \otimes b = ba'$ (in particular, $aa' = a' \otimes b$).*
8. *If A and B are square matrices of respective size m and n,*

$$\det\left(A \otimes B\right) = \left(\det A\right)^m \left(\det B\right)^n.$$

9. $Tr\left(A \otimes B\right) = Tr\left(A\right) Tr\left(B\right)$.

10. $rank\,(A \otimes B) = rank\,(A)\,rank\,(B)$*.*

11. The singular values of $A \otimes B$ *are the product of the singular values of* A *and those of* B*.*

Let f be a smooth map from the space of $m \times p$ matrices to the space of $n \times q$ matrix. Define $\frac{df(A)}{dA}$ as the $(nq) \times (mp)$ matrix such that for an $m \times p$ matrix X,

$$vec\left(\lim_{e \to 0} \frac{f(A + eX) - f(A)}{e}\right) = \frac{df(A)}{dA}.vec(X).$$

We use the notation A^{-*} for $(A^*)^{-1}$.

Fact 3. *Let* A *be a* $m \times p$ *matrix and* B *an* $n \times q$ *matrix. Then:*

1. $\frac{d(AXB)}{dX} = B^* \otimes A$*.*

2. $\frac{dA^*}{dA} = \mathbb{T}_{m,p}$*.*

3. $\frac{dA^{-1}}{dA} = -\left(A^{-*} \otimes A^{-1}\right)$*.*

4. $\frac{dA^2}{dA} = I \otimes A + A^T \otimes I$

5. For A *symmetric,* $\frac{dA^{1/2}}{dA} = \left(I \otimes A^{1/2} + A^{1/2} \otimes I\right)^{-1}$*.*

§ *École polytechnique, Department of Economics. Address: Route de Saclay, 91128 Palaiseau, France. E-mail: raicho.bojilov@polytechnique.edu.*

† *Sciences Po, Department of Economics. Address: 27 rue Saint-Guillaume, 75007 Paris, France. E-mail: alfred.galichon@sciences-po.fr.*

References

[1] Becker, G. (1973). "A theory of marriage, part I," *Journal of Political Economy*, 81, pp 813-846.

[2] Ben-Akiva, M. and Watanatada, T. (1981). "Application of a Continuous Spatial Choice Logit Model." In *Structural Analysis of Discrete Choice Data with Econometric Applications* (C.F. Manski and D. McFadden, eds.), MIT Press.

[3] Ben-Israel, Adi; Thomas N.E. Greville (2003). *Generalized Inverses*. Springer-Verlag.

[4] Chiappori, P.-A., Oreffice, S. and Quintana-Domeque, C. (2012). "Fatter attraction: anthropometric and socioeconomic matching on the marriage market," *Journal of Political Economy* 120 (4), pp. 659–695.

[5] Chiappori, P.-A., Salanié, B., and Weiss, Y. (2008). "Assortative matching on the marriage market: a structural investigation," working paper.

[6] Choo, E., and Siow, E. (2006). "Who marries whom and why". *Journal of Political Economy* 114, pp. 172–201.

[7] Cosslett, S. (1988). "Extreme-value stochastic processes: A model of random utility maximization for a continuous choice set," Technical report, Ohio State University.

[8] Dagsvik, J. (1994). "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes," *Econometrica* 62, 1179–1205.

[9] Decker, C., Lieb, E., McCann, R. and Stephens, B. (2013). "Unique equilibria and substitution effects in a stochastic model of the marriage market". *Journal of Economic Theory* 148, pp. 778–792.

[10] Dowson, D.C., and B.V. Landau (1982). "The Fréchet distance between multivariate normal distributions". *Journal of Multivariate Analysis* 12, pp. 450–455.

[11] Dupuy, A., and Galichon, A. (2014). "Personality traits and the marriage market". Forthcoming in the *Journal of Political Economy.*

[12] Fox, J. (2010). "Identification in Matching Games". *Quantitative Economics* 1(2), pp. 203–254.

[13] Fox, J. (2014). "Estimating Matching Games with Transfers". Working paper.

[14] Gabaix, X. and Landier (2008). "Why Has CEO Pay Increased So Much?". *Quaterly Journal of Economics* 123(1), p. 49–100.

[15] Galichon, A., and Salanié, B. (2010). "Matching with Trade-offs: Revealed Preferences over Competing Characteristics". Technical report.

[16] Galichon, A., and Salanié, B. (2014). "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models". Working paper.

[17] Graham, B. (2011). "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. Jackson. Elsevier.

[18] Hsieh, C.-T., Hurst, E., Jones, C., Klenow, P. (2013). "The Allocation of Talent and US Economic Growth." *NBER Working Paper* No. 18693.

[19] Lindenlaub I. (2013). "Sorting Multidimensional Types: Theory and Application" Manuscript.

[20] Menzel, K. (2014). "Large Matching Markets as Two-Sided Demand Systems." Manuscript.

[21] Olkin, I. and F. Pukelsheim (1982). "The distance between two random vectors with given dispersion matrices." *Linear Algebra Appl.* 48, 257-263.

[22] Resnick, S., and R. Roy (1991). "Random USC functions, Max-stable processes and continuous choice," *Annals of Applied Probability* 1(2), 267–292.

[23] Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers* 3 (2), 135–146.

[24] Tervio, M. (2008). "The difference that CEO make: An Assignment Model Approach," *American Economic Review* 98 (3), pp. 642–668.

[25] Villani, C. (2004). *Topics in Optimal Transportation.* AMS.